Università di Bologna

Recent advances in approximation using Krylov subspaces

V. Simoncini

Dipartimento di Matematica, Università di Bologna
and CIRSA, Ravenna, Italy
valeria@dm.unibo.it

Mostly work in collaboration with D. Szyld, Temple University

The framework

It is given the problem: $\quad$ *Find* vector x : $G(\mathcal{A}, x) = 0$

The operator $v \mapsto \mathcal{A}(v)$ is known thru an approximation:

$$v \mapsto \mathcal{A}_\epsilon(v) \quad \text{where} \quad \mathcal{A}_\epsilon \to \mathcal{A} \ \text{ for } \ \epsilon \to 0$$

(ϵ may be tuned)

Efficiently approximate x in the space

$$\mathcal{K}_m = \text{span}\{v, \mathcal{A}_{\epsilon_1}(v), \mathcal{A}_{\epsilon_2}(\mathcal{A}_{\epsilon_1}(v)), \ldots\}, \quad v \in \mathbb{C}^n$$

with $\dim(\mathcal{K}_m) = m$

⋆ for $\mathcal{A} = A$, $\epsilon = 0$ $\quad \Rightarrow \quad$ $\mathcal{K}_m = \text{span}\{v, Av, A^2 v, \ldots, A^{m-1} v\}$

Examples of $\mathcal{A}$:

- Solution of (preconditioned) large linear systems,

$$Ax = b \qquad n \times n \qquad \mathcal{A} = A$$

(also Schur complements, multipole, etc)

- Shift-and invert procedures for interior eigenvalues

$$Ax = \lambda Mx, \qquad \|x\| = 1, \qquad \mathcal{A} = (\sigma M - A)^{-1}$$

- Rational approximation. e.g. preconditioned exponential approximation

$$x = \exp(A)v, \qquad \mathcal{A} = (\gamma I - A)^{-1}$$

- ...

The exact approach

To focus our attention: $\mathcal{A}(v) = \mathcal{A}v$

$\mathcal{K}_m(\mathcal{A}, v)$ Krylov subspace $\qquad V_m$ orthogonal basis

Key relation in Krylov subspace methods:

$$AV_m = V_{m+1}\underline{H}_m \qquad v = V_{m+1}e_1\beta \qquad \underline{H}_m = \begin{bmatrix} H_m \\ h_{m+1,m}e_m^T \end{bmatrix}$$

$$x_m \in \mathcal{K}_m \quad \Rightarrow \quad x_m = V_m y_m \quad \Rightarrow \quad r_m = G(\mathcal{A}, x_m)$$

$$\mathcal{A} \to \mathcal{A}_\epsilon \quad \Rightarrow \quad r_m = G(\mathcal{A}_\epsilon, x_m)$$

The inexact key relation

$$\mathcal{A}_{\epsilon_j} v = \mathcal{A} v + f_j \qquad \|f_j\| = O(\epsilon_j), \quad j = 1, 2, \ldots$$

$$\mathcal{A} V_m = V_{m+1} \underline{H}_m + \underbrace{F_m}_{[f_1, f_2, \ldots, f_m]} \qquad F_m \text{ error matrix}$$

How large is F_m allowed to be?

Claim: the perturbation induced by ϵ_j may be far less devastating for $x_m \rightarrow x$ than $|\epsilon_j|$ would predict

$$\mathcal{A} x_m = \mathcal{A} V_m y_m = V_{m+1} \underline{H}_m y_m + F_m y_m$$

$$\|F_m y_m\| \quad \text{small then} \quad V_{m+1} \underline{H}_m y_m \approx \mathcal{A} x_m$$

A dynamic setting

$$F_m y = [f_1, f_2, \ldots, f_m] \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_m \end{bmatrix} = \sum_{i=1}^{m} f_i \eta_i$$

◇ The terms $f_i \eta_i$ need to be small:

$$\|f_i \eta_i\| < \frac{1}{m}\epsilon \quad \forall i \quad \Rightarrow \quad \|F_m y\| < \epsilon$$

◇ If $|\eta_i|$ small $\quad \Rightarrow \quad \|f_i\|$ is allowed to be large

⋆ In several problems it can be shown that $\qquad |\eta_i| \leq \gamma_m \|r_{i-1}\|$

A refinement procedure

Example: linear system $Ax = b$ with residual minimizing method

$r_k = b - Ax_k, \quad r_m = b - Ax_m$ residuals after k, m its, $\quad k < m$

$$\|V_m y_m - V_k y_k\| \leq 2\|A^{-1}\| \, \|r_k\|$$

$\left(\|V_m y_m - V_k y_k\| \leq \|A^{-1}\| \, \|AV_m y_m - AV_k y_k\| = \|A^{-1}\| \, \|r_m - r_k\| \leq 2\|A^{-1}\| \, \|r_k\| \right)$

In general, in $G(\mathcal{A}, V_m y_m) = 0$ we expect a relation of the type

$$y_m = \begin{bmatrix} y_k \\ 0 \end{bmatrix} + d_m, \qquad \|d_m\| \leq \gamma_m \|r_k\|$$

The correction term

$$y_m = \begin{bmatrix} y_k \\ 0 \end{bmatrix} + d_m, \qquad \|d_m\| \le \gamma_m \|r_k\|$$

Explicitly write d_m so that:

- Linear system. GMRES
  (Simoncini & Szyld '03, Sleijpen & van den Eshof '04)

$$\gamma_m = \frac{1}{\sigma_{\min}(\underline{H_m})}$$

- Linear system. FOM
  (Simoncini & Szyld '03, Sleijpen & van den Eshof '04)

$$d_m = -H_m^{-1} \begin{bmatrix} 0 \\ e_1 \end{bmatrix} \|r_k\|, \quad \gamma_m = \|H_m^{-1}\|$$

- Eigenproblem. Ritz value θ_k of H_k (under suitable hypotheses) (Simoncini, SINUM To appear)

$$\gamma_m = \frac{2}{\delta_{m,k}}, \qquad \delta_{m,k} = \sigma_{\min}(H_m - \theta_k I)$$

  Similar results for Harmonic Ritz values and Lanczos approximations

Note: relations also hold for problem $G(\mathcal{A}_\epsilon, V_m y_m) = 0$!!

see Simoncini & Szyld, GAMM Proceedings '05

Relaxing the accuracy in linear systems

$A \cdot v_i$ not performed exactly $\quad \Rightarrow \quad (A + E_i)v_i = Av_i + f_i$

$$b - Ax_m = V_{m+1}(e_1\beta - \underline{H}_m y_m) - F_m y_m$$

For instance, for GMRES: If $\quad \|E_i\| \le \frac{\sigma_{\min}(\underline{H}_m)}{m} \frac{1}{\|r_{i-1}\|}\varepsilon \quad i = 1, \ldots, m$

then $\quad \|F_m y_m\| \le \sum_{i=1}^{m} \|E_i\| |\eta_i| \le \varepsilon \quad$ so that

$$\|(b - Ax_m) - V_{m+1}(e_1\beta - \underline{H}_m y_m)\| \le \varepsilon$$

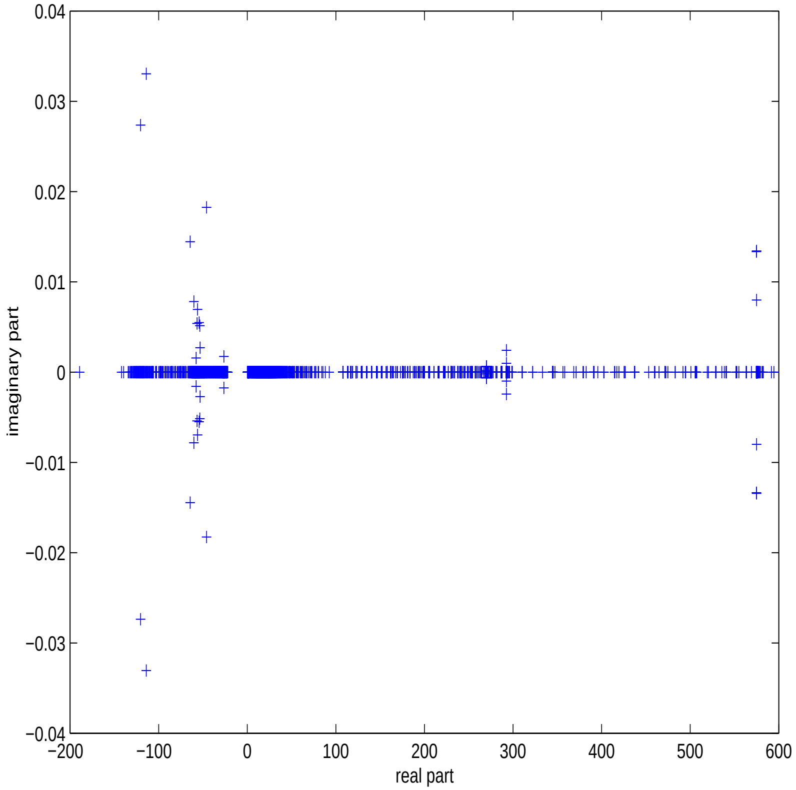
(see also Bouras & Frayssè '05, Giraud & Gratton & Langou, tr. '04)

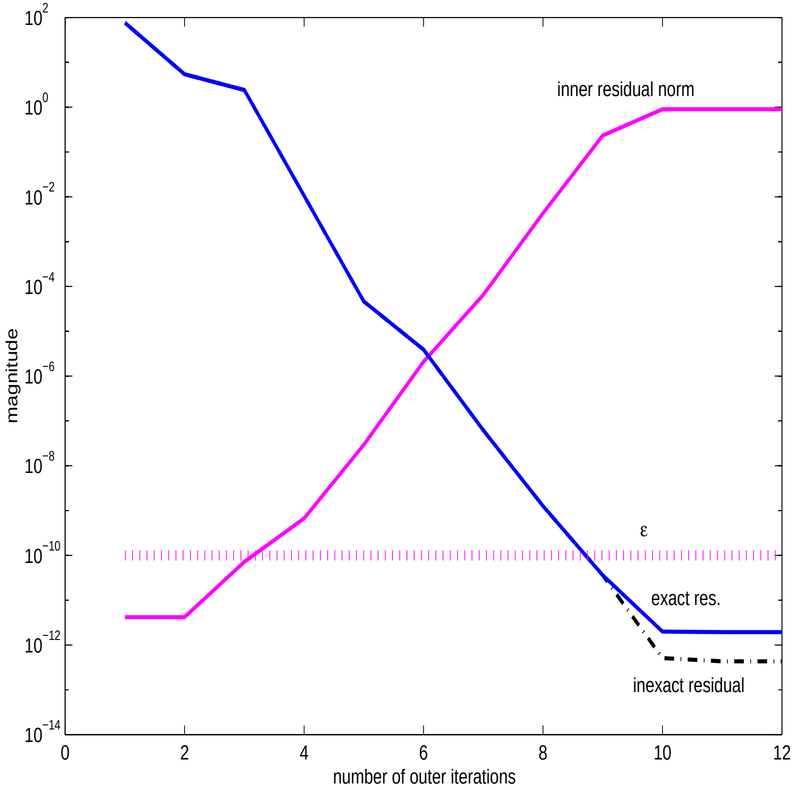
Note: $\quad \|b - Ax_m\| \le \varepsilon \quad$ final attainable residual norm

Eigenproblem

Inverted Arnoldi: $Ax = \lambda x$ Find $\min |\lambda|$ $y \leftarrow \mathcal{A}(v) = A^{-1}v$

Matrix SHERMAN5





Rational approximation in $\mathcal{K}_m$

$$\Psi_\nu(A)x = \Phi_\nu(A)v, \qquad \Psi_\nu, \Phi_\nu \in \mathbb{P}_\nu$$

e.g. $x \approx \exp(A)v$

Approximate x in $\mathcal{K}_m(A, v)$

Write $\quad \Psi_\nu(\lambda)^{-1}\Phi_\nu(\lambda) = \tau_0 + \sum_{k=1}^{\nu} \tau_k(\lambda - \lambda_k)^{-1}$

Therefore,

$$x = \Psi_\nu(A)^{-1}\Phi_\nu(A)v = \tau_0 v + \sum_{k=1}^{\nu} \tau_k(A - \lambda_k I)^{-1} v$$

$$x = \Psi_\nu(A)^{-1}\Phi_\nu(A)v = \tau_0 v + \sum_{k=1}^{\nu} \tau_k (A - \lambda_k I)^{-1} v$$

Krylov subspace approximation:

$$x_m = \tau_0 v + \sum_{k=1}^{\nu} \tau_k V_m y_m^{(k)}$$

$\Rightarrow$ $y_m^{(k)}$'s have decreasing pattern (see Lopez & Simoncini, tr.'05)

$$\Downarrow$$

Possible use in preconditioning techniques
(e.g. van den Eshof & Hochbruck, tr.'04)

Inexactness when A symmetric

A symmetric $\quad \Rightarrow \quad A + E_i$ nonsymmetric

- Assume $V_m^T V_m = I \quad \rightarrow \quad H_m$ upper Hessenberg
- Wise implementation of short-term recurr. /truncated methods (V_m non-orth. $\rightarrow W_m, \qquad H_m$ tridiag./banded $\rightarrow T_m$)

  - Inexact short-term recurrence system solvers

  (Golub-Overton '88, Golub-Ye '99, Notay '00, Sleijpen-van den Eshof '04, ...)

  - Inexact symmetric eigensolvers

  (Lai-Lin-Lin 1997, Golub-Ye 2000, Golub-Zhang-Zha 2000, Notay 2002, ...)

  - Truncated methods (Simoncini & Szyld, Num. Math. to appear)

References

1. V. Simoncini and D. B. Szyld. Theory of Inexact Krylov Subspace Methods and Applications to Scientific Computing. *SIAM J. Scientific Comput.* v.25, n.2 (2003), pp. 454-477.

2. F. Perotti and V. Simoncini. Analytical and Numerical Techniques in frequency domain response computation. *Recent Research Developments in Structural Dynamics*, A. Luongo Ed. Research Signpost Pub., 2003, pp.33-54.

3. V. Simoncini and D. B. Szyld. On the occurrence of Superlinear Convergence of Exact and Inexact Krylov Subspace Methods. *SIAM Review 2005*.

4. V. Simoncini. Variable accuracy of matrix-vector products in projection methods for eigencomputation. *SINUM, To appear.*

5. V. Simoncini and D. B. Szyld. The effect of non-optimal bases on the convergence of Krylov Subspace Methods. *Num. Math. To appear.*

6. L. Lopez and V. Simoncini. Analysis of projection-type methods for approximating the matrix exponential operator. March 2005.

More at `www.dm.unibo.it/~simoncin`